

On the optimal allocation of students when peer effects are at work: tracking vs. mixing

Marisa Hidalgo-Hidalgo

Received: 28 May 2009 / Accepted: 5 May 2010 / Published online: 26 May 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract The belief that both the behavior and outcomes of students are affected by their peers is important in shaping education policy. I analyze two polar education systems -tracking and mixing- and propose several criteria for their comparison. I find that tracking is the system that maximizes average human capital. Nevertheless, no single system is unanimously preferred by the entire population. In addition, I find that the gains of tracking, compared to mixing, are sensitive to changes in the level of dispersion in the pre-school achievement distribution of the population.

Keywords Human capital · Efficiency · Peer effects · Tracking · Mixing

JEL Classification D63 · I28 · J24

This paper is part of my Ph.D. dissertation, completed at Universidad de Alicante under the very helpful supervision of Iñigo Iturbe-Ormaetxe. I am grateful for the comments and suggestions of Gianni De Fraja, Elena Del Rey, Moshe Justman, Francois Maniquet, Francisco Martínez-Mora and Ignacio Ortuño-Ortín. Thanks are due to the editor, María-Angeles de Frutos, and two anonymous referees for valuable comments that substantially improved the analysis of this paper. I also acknowledge members of the Department of Economics at University of Leicester, where part of this work was done, for their hospitality. Financial support from the Spanish Ministry of Education and Science (SEJ 2007-67734), Junta de Andalucía (SEJ-2905) and the Instituto Valenciano de Investigaciones Económicas is gratefully acknowledged. All remaining errors are mine.

M. Hidalgo-Hidalgo (✉)

Departamento de Economía, Métodos Cuantitativos e Historia Económica, Universidad Pablo de Olavide de Sevilla, Ctra. de Utrera, Km. 1, 41013 Sevilla, Spain
e-mail: mhidalgo@upo.es

1 Introduction

Peer effects are at the heart of many recent debates on educational reforms.¹ The critical importance to both parents and policy makers of peer group distribution in school is indisputable. Given the existence of peer effects, defined here as the effect of the ability distribution of her peers on an individual's academic performance, governments should keep peer effects in mind when planning how to best meet their educational policy objectives. One situation in which peer effects must be carefully considered is when governments choose whether to stream (track) or to mix students of differing abilities within public schools. Under a tracking system, schools are hierarchically organized to accommodate a range of student performance levels, and students are placed in the school that best suits their ability level. By contrast, mixing works by grouping students of differing abilities within the same schools.

The peer group quality affects student achievement positively.² However, raising peer quality for every student is an impossible task. From a policy point of view, the more relevant questions are concerned with efficiency issues. This paper analyzes efficiency, among other criteria, when comparing tracking and mixing systems. It contributes to the comparison by addressing three main questions. First, it analyzes which system is chosen by the population under a majority voting rule. Second, it asks which system maximizes average human capital at the compulsory level and considers how the degree of dispersion in the pre-school achievement distribution of the population affects this issue. Finally, it explores whether or not the entire population can be said to prefer one of the aforementioned systems -tracking or mixing.

To address these issues I introduce a model in which students differ in parental background as well as in pre-school achievement, which will be positively correlated. The production of human capital depends on both students' previous achievements and peer group characteristics. The degree of complementarity between both inputs, together with the degree of dispersion in pre-school achievement, is shown to be critical for the comparison between tracking and mixing systems.

The contribution of my paper to the relevant literature is twofold. First, it analyzes the role of dispersion in pre-school achievement on the optimal allocation of students. In addition, assuming concavity in peer effects, I discuss how the complementarity between students' previous achievements and peer characteristics can determine which system maximizes average human capital. Observe that there is a clear trade-off between both properties: while concavity implies an efficiency loss from tracking, the higher the degree of complementarity, the higher the efficiency gains from tracking. Second, the paper derives the induced distribution of achievements under each educational system and compares them, according to several criteria. This is important, as in the comparison of the two systems the previous literature has only focused on mean

¹ See, for example, the "NBER Reporter Online: Fall 2006" available at <http://www.nber.org/reporter/fall06/#report>.

² There is a large empirical literature on peer effects on individuals' educational achievements. See Hoxby (2000), Ding and Lehrer (2007) and Kang (2007), among others, who find evidence of significant peer effects on achievement.

effects (see [Arnott and Rowse 1987](#); [Benabou 1996](#), among others).³ In particular, the “veil of ignorance” approach will be used.

First, my study suggests that, when choosing the education system by majority voting, a non-elitist tracking system, i.e., a system in which middle class households get their kids into the high track, is more likely to be politically supported in a democracy than an elitist tracking system that leaves them out. This is due to the positive correlation between the probability of being assigned to the high track and family background, which is supported by the empirical literature. Second, I find that tracking is the system that maximizes average human capital at the end of the compulsory level. In addition, it is shown that the higher the complementarity between peer effects and individual pre-school achievement, the larger the difference between average human capital under tracking and mixing. Interestingly, I find that the level of dispersion in the distribution of pre-school achievements plays a key role when comparing average human capital under tracking and mixing. For example, focus first on societies in which the distribution of pre-school achievement is not very dispersed. In this scenario, as some dispersion appears, tracking becomes more and more attractive than mixing for maximizing average human capital. In this case, the complementarity between the peer group effect and the pre-school achievement level offsets the concavity effect and drives the result. The gains of the high-achievers (as compared to the mixing case) more than compensate for the losses of the low-achievers. However, this is not the case in societies in which the pre-school achievement distribution is much dispersed. In this scenario, additional dispersion makes tracking relatively less attractive than mixing for maximizing average human capital. Clearly, complementarity between the peer group effect and the pre-school achievement level still offsets the concavity effect, but the losses of the low-achievers might be balanced with the gains of the high-achievers (compared to mixing). Finally, my study suggests that among all risk-averse households with uncertainty about their kids’ pre-school achievements, there is no unanimity on the choice of the educational system.

The rest of the paper is organized as follows. Section 2 describes the model and the main features of human capital distribution under the two education systems at the compulsory school level. Section 3 compares the induced distributions of human capital in these two systems. Section 4 concludes.

2 Model

2.1 Individuals

Consider a population of size 1. Individuals differ in two aspects: their family background and their pre-school achievement, θ_0 , where $\theta_0 \in [0, 1]$. To make the model tractable, I assume that there are only two values for family background, that is,

³ [Arnott and Rowse \(1987\)](#) analyze the optimal allocation of students and resources by focusing on the degree of concavity of the peer group effect. [Benabou \(1996\)](#) studies the general trade-off between complementarity and curvature when studying optimal community composition.

individuals can have either poor or rich parents, with probabilities $1 - \lambda$ and λ , respectively.⁴ I denote by $g_b(\theta_0)$ the p.d.f. (probability density function) of θ_0 , conditional on having family background b , where $b = p, r$ for poor and rich parents, respectively. Richer individuals have more resources to invest in their kids. In addition, richer individuals are more educated and care more about education, which positively influences their children's achievement levels upon entering school. Thus to capture the possibility that some level of positive dependence exists between parental background and pre-school achievement, I assume that $g_p(\theta_0) = \gamma\theta_0^{\gamma-1}$ and $g_r(\theta_0) = 1$, where $\gamma \in (0, 1]$ measures the pre-school achievement gap between poor and rich students. Thus the cumulative distribution function (CDF) of pre-school achievement, denoted by $G(\theta_0)$, can be expressed as:⁵

$$G(\theta_0) = \begin{cases} 0 & \text{if } \theta_0 < 0 \\ (1 - \lambda)\theta_0^\gamma + \lambda\theta_0 & \text{if } 0 \leq \theta_0 \leq 1 \\ 1 & \text{if } \theta_0 > 1. \end{cases} \quad (1)$$

Figure 1 illustrates the CDF of θ_0 in Eq. (1) for two societies, A and B, that differ in the pre-school achievement gap between poor and rich students, γ , but share the same proportion of rich individuals, λ . This difference may stem from the fact that A spends more resources in early childhood education than B does. In particular, I set $\gamma = 0.8$ and $\gamma = 0.2$ for societies A and B, respectively, and $\lambda = 5/6$.⁶

From Eq. (1) we can check that $G(\theta_0)$ is decreasing in γ . This is illustrated also in Fig. 1. That is, provided that the proportion of rich individuals is the same in both societies, because in society A the parameter γ is higher than in society B, the distribution of preschool achievement in B is dominated by that in A. In other words, the probability of having a high (low) level of pre-school achievement in society A is higher (lower) than in society B. In addition, note from Eq. (1) that $G(\theta_0)$ is decreasing in the proportion of rich individuals, λ .

Observe that because the gap in pre-school achievement between poor and rich people is lower in society A than in B, not only the pre-school achievement distribution is more dispersed in B than in A, but mean pre-school achievement is higher in A than in B. Therefore, it is interesting to analyze the coefficient of variation of θ_0 ,

⁴ Alternatively we could interpret the two parent types as white or black, natives or immigrants, etc.

⁵ See Sect. 3.2 for an analysis of the robustness of results with respect to other pre-school achievement distribution specifications.

⁶ This means that the mean pre-school achievement level for poor students represents 89% and 33% of the achievement level for rich students in societies A and B, respectively. According to the PISA 2003 results for OECD countries, the mean in math performance among immigrants represents about 90%, on average, of the mean in math performance for native students. The value for λ is drawn from OECD data. In most studies on poverty, the poor are all individuals living in households with incomes below the poverty line, which is fixed at 50% of the national median adjusted income (see Collado and Iturbe-Ormaetxe 2008). According to the OECD questionnaire on household income distribution (2002), those individuals whose equivalent disposable income in 2000 was less than 50% of the population median averaged 10%, with an increasing tendency (see OECD 2002). Hence $\lambda = 5/6$ seems appropriate if we, alternatively, interpret λ as the proportion of natives in the population.

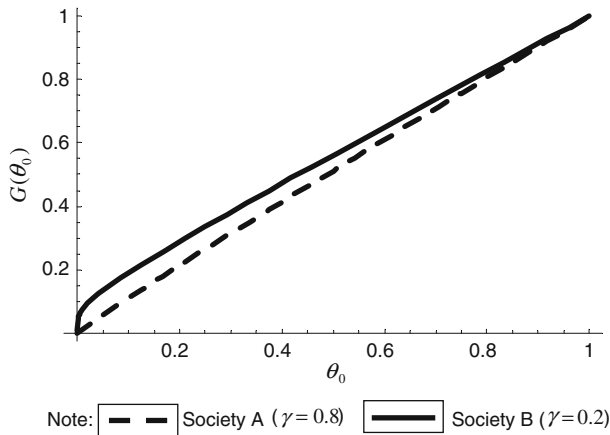


Fig. 1 The distribution of pre-school achievement

$cv_{\theta_0}(\gamma, \lambda)$. From (1) it is:

$$cv_{\theta_0}(\gamma, \lambda) = \frac{2(1+\gamma)}{2\gamma + \lambda(1-\gamma)} \sqrt{\frac{\lambda}{3} + \frac{\gamma(1-\lambda)}{2+\gamma} - \frac{(2\gamma + \lambda(1-\gamma))^2}{4(1+\gamma)^2}}, \quad (2)$$

which is strictly decreasing in γ and λ .⁷ That is, the degree of dispersion in the distribution of pre-school achievement is controlled by γ (the higher γ , the lower the gap between rich and poor and hence the less dispersion) and by λ (the higher λ , the less proportion of poor families and hence the less dispersion). Below we analyze the role of the degree of dispersion in pre-school achievement on the comparison between tracking and mixing systems, by examining the effects of changes in γ and/or λ , instead of examining the effect of changes in the coefficient of variation.

Finally, observe that the distribution of pre-school achievement θ_0 is right-skewed, that is, the median value of the distribution of θ_0 is lower than its average for any $\lambda \in (0, 1)$ and $\gamma \in (0, 1]$. The skewness of the pre-school achievement distribution might be a controversial assumption. Note that, due to the existing positive dependence between pre-school achievement and parental income, the distribution of both variables might have a similar shape (see Becker 1993, among others). In addition, most empirical evidence shows that income distribution is right-skewed (for example, in 1989, mean and median US household incomes were \$36,250 and \$28,906). While it is not obvious that this is true in all cases, I find it to be a plausible assumption. It is in fact the most usual assumption in the theoretical literature regarding both earnings distribution (see, for example, Epple and Romano 1998) and pre-school

⁷ It can be checked that $\frac{\partial cv_{\theta_0}(\gamma, \lambda)}{\partial \gamma} = -\frac{2[\lambda(1-\gamma)(5+\gamma)+6\gamma](1-\lambda)(\gamma+1)^2}{3(\gamma^2+3\gamma+2)^2\sigma}$ and that $\frac{\partial cv_{\theta_0}(\gamma, \lambda)}{\partial \lambda} = -\frac{2[\lambda(1-\gamma)+\gamma](1-\gamma)(\gamma+1)^2}{(\gamma^2+3\gamma+2)(\lambda+2\gamma-\lambda\gamma)^2\sigma}$ where the standard deviation of θ_0 is given by $\sigma = \sqrt{\frac{\lambda}{3} + \frac{\gamma(1-\lambda)}{2+\gamma} - \frac{(2\gamma + \lambda(1-\gamma))^2}{(2+2\gamma)^2}}$.

achievement distribution (Gradstein and Justman 2004). See Sect. 3.2 for an analysis of the sensitivity of the results of the paper to this assumption.

Individuals accumulate human capital by attending compulsory education, which is free of charge, and they are not allowed to work.

2.2 Production of human capital

At the compulsory level, individuals are separated into different groups or classes according to their pre-school achievement levels. To simplify, I consider only two groups. The production of human capital depends on two factors. The first is the individual's pre-school achievement, θ_0 .⁸ The second is the “peer group” effect that depends on the characteristics of the group in which the individual is placed. These characteristics are summarized by the mean achievement of the group j , or the “peer” effect, denoted by $\bar{\theta}_0^j$. After attending compulsory education, an individual with pre-school achievement θ_0 ends up with a level of human capital θ_1 .⁹

I assume that the production of human capital θ_1 is a twice-differentiable, increasing and concave function. In particular, I assume it follows a CES production function with inputs θ_0 and $\bar{\theta}_0^j$.¹⁰

$$\theta_1(\theta_0, \bar{\theta}_0^j) = A \left(\rho \theta_0^\beta + (1 - \rho) (\bar{\theta}_0^j)^\beta \right)^{\frac{1}{\beta}}, \quad (3)$$

where $A > 1$, $\rho \in [0, 1]$ and $\beta \in (0, 1]$. The parameter ρ captures the weight of pre-school achievement on θ_1 . The parameter β determines the elasticity of substitution between the two inputs $\bar{\theta}_0^j$ and θ_0 . First, regarding pre-school achievement, the individuals' levels of human capital are an increasing function of previous achievement levels, but at a decreasing rate. Regarding the peer group effect, and according to the empirical evidence, we observe that it is non-linear: the individual's human capital rises with an improvement in the average pre-school achievement within the classroom, but this positive effect has decreasing returns.¹¹ The empirical evidence

⁸ Observe that both of the individual's characteristics, pre-school achievement and parental background, affect her final human capital, but in a different way. Whereas an individual's pre-school achievement has a *direct* effect on it (because further human capital builds on previous achievement), parental background has an *indirect* effect through the positive dependence on an individual's pre-school achievement.

⁹ There is a large empirical literature on peer effects and still an open debate on the influence of peers on individual educational achievements, with some studies finding little or no effects (e.g., Angrist and Lang 2004; Foster 2006) and others finding large effects (e.g., Hoxby and Weingarth 2006; Ding and Lehrer 2007). However, this assumption is commonly accepted in related literature. See, among others, Epplé and Romano (1998), Epplé et al. (2002) and Bishop (2006), who also assume that peers affect an individual through the mean of their characteristics.

¹⁰ See Sect. 3.3 for an analysis of the robustness of results with respect to this human capital production function specification.

¹¹ The importance of the individual's prior achievement levels on the individual's acquisition of human capital has been explored theoretically as well as empirically (see, for example, Heckman 2006). Ding and Lehrer (2007) and Hoxby and Weingarth (2006), among others, suggest that the peer group effect is non-linear.

regarding the relationship between individuals' abilities and the peer group effect is still mixed. Hoxby and Weingarth (2006) cannot reject the hypothesis according to which high and low-achieving students benefit equally from the presence of high-achieving students. However, Ding and Lehrer (2007) find that high-ability students benefit more from having higher-achieving schoolmates than students of lower ability. Hence the functional form of Eq. (3) allows for studying how the complementarity between peer effects and individuals' abilities affects the comparison between tracking and mixing systems.

Therefore, on the one hand, the positive impact of an increase in mean achievement is decreasing, resulting in an efficiency loss from tracking or streaming students of differing abilities within groups. On the other hand, Eq. (3) allows for the possibility that $\bar{\theta}_0^j$ and θ_0 are either complements or substitutes. Note here that for any $\beta < 1$, we have that $\frac{\partial^2 \theta_1}{\partial \theta_0 \partial \bar{\theta}_0^j} > 0$, that is, high-achievers benefit most from an increase in mean achievement, which implies that there would be an efficiency gain from streaming students.¹² Therefore, Eq. (3) clearly sets up a tension between mixing and tracking.¹³

2.3 Education systems at compulsory level

In this section I describe the two polar education systems of mixing and tracking and analyze the distribution of human capital at the end of compulsory school under each system.

2.3.1 Mixing

In a mixing system, the pre-school achievement distribution is the same in both classrooms. The average pre-school achievement within each classroom, denoted by $\bar{\theta}_0^m(\gamma, \lambda)$, coincides with the average pre-school achievement in the population:

$$\bar{\theta}_0^m(\gamma, \lambda) = (1 - \lambda) \left(\frac{\gamma}{\gamma + 1} \right) + \frac{\lambda}{2}. \quad (4)$$

Observe that $\bar{\theta}_0^m(\gamma, \lambda) \leq 1/2$ for any γ and λ . In a mixing system, θ_1 will lie in the support $[\underline{m}, \bar{m}]$, where \underline{m} and \bar{m} stand for $\underline{m}(\gamma, \lambda)$ and $\bar{m}(\gamma, \lambda)$ and denote the level of human capital θ_1 acquired in a mixing system by the “worst” (lowest pre-school achiever) and the “best” (highest pre-school achiever) individuals in the population,

¹² In particular, for β close to 0, both $\bar{\theta}_0^j$ and θ_0 have some level of complementarity and as β tends to 1, the two factors become perfect substitutes.

¹³ See also Benabou (1996) who, in addition to these two effects, considers whether peer effects are stronger in classrooms with a large or small share of rich individuals. Note that this last effect cannot be analyzed here because, as we will see below, the way peer effects depend on the proportion of rich individuals varies with the education system and the classroom type.

respectively:

$$\underline{m}(\gamma, \lambda) = A(1 - \rho)^{\frac{1}{\beta}} (\bar{\theta}_0^m)^{\frac{1}{\beta}}. \quad (5)$$

$$\bar{m}(\gamma, \lambda) = A \left(\rho + (1 - \rho) (\bar{\theta}_0^m)^{\beta} \right)^{\frac{1}{\beta}}. \quad (6)$$

Therefore, from Eq. (3) the CDF of θ_1 in a mixing system, denoted $F_M(\theta_1)$, is:

$$F_M(\theta_1) = \begin{cases} 0 & \text{if } \theta_1 < \underline{m} \\ (1 - \lambda)\varphi^m(\theta_1)^\gamma + \lambda\varphi^m(\theta_1) & \text{if } \underline{m} \leq \theta_1 \leq \bar{m} \\ 1 & \text{if } \theta_1 > \bar{m}, \end{cases} \quad (7)$$

where $\varphi^m(\theta_1) = \left(\frac{1}{\rho} \left(\left(\frac{\theta_1}{A} \right)^\beta - (1 - \rho)(\bar{\theta}_0^m)^\beta \right) \right)^{\frac{1}{\beta}}$. It is straightforward to check that $\partial F_M(\theta_1)/\partial \gamma$ is negative and hence $F_M(\theta_1)$ is decreasing in γ . This is because an increase in γ makes the distribution of pre-school achievement closer to that of rich individuals. Observe from Eq. (4) that average pre-school achievement $\bar{\theta}_0^m(\gamma, \lambda)$ (which is the only determinant of the difference in human capital between both societies) is increasing in γ . I denote by $E_M(\theta_1)$ the expected value of θ_1 in a mixing system, where:

$$E_M(\theta_1) = \int_{\underline{m}}^{\bar{m}} \theta_1 f_M(\theta_1) d\theta_1 = \bar{m} - \int_{\underline{m}}^{\bar{m}} F_M(\theta_1) d\theta_1, \quad (8)$$

and $f_M(\theta_1)$ denotes the p.d.f. of θ_1 in a mixing system. From Eqs. (7) and (8), it is immediate that the average human capital in a mixing system increases as the dispersion level in pre-school achievement distribution decreases (measured by an increase in either γ or λ). Again, this is due to the fact that the mean pre-school achievement in the population is increasing in γ and λ .

2.3.2 Tracking

Tracking students implies grouping them on the basis of pre-school achievement. For the sake of simplicity I permit only two tracks and I denote by δ and $(1 - \delta)$ the proportion of students in the low and the high tracks, respectively. In addition, I denote by t the threshold level of pre-school achievement used for grouping students into one track or the other. Thus a student is assigned to the high (low) track when her pre-school achievement θ_0 is above (below) t . From (1), the threshold level t is implicitly defined as:

$$(1 - \lambda)t^\gamma + \lambda t = \delta. \quad (9)$$

Computing the corresponding derivatives in (9), it can be checked that the threshold t is increasing in λ and γ . Recall that the distribution of pre-school achievement θ_0 is right-skewed. This implies that the proportion of rich students among all high-track students (i.e., $\frac{\lambda(1-t)}{(1-\delta)}$) will be higher than the proportion of rich students in the population, λ . This captures the empirical evidence found by Brunello and Checchi (2007), among others, regarding the socioeconomic composition of the different tracks.

I denote by $\bar{\theta}_0^l(\gamma, \lambda, \delta)$ and $\bar{\theta}_0^h(\gamma, \lambda, \delta)$ the average pre-school achievement in the low and high tracks, respectively. Thus given the distributional assumptions on θ_0 :

$$\bar{\theta}_0^l(\gamma, \lambda, \delta) = \frac{\int_0^t \theta_0 g(\theta_0) d\theta_0}{\int_0^t g(\theta_0) d\theta_0} = \frac{\left(\frac{\gamma}{\gamma+1}\right) (1-\lambda)t^{\gamma+1} + \frac{1}{2}\lambda t^2}{\delta}, \quad (10)$$

and:

$$\bar{\theta}_0^h(\gamma, \lambda, \delta) = \frac{\int_t^1 \theta_0 g(\theta_0) d\theta_0}{\int_t^1 g(\theta_0) d\theta_0} = \frac{\left(\frac{\gamma}{\gamma+1}\right) (1-\lambda) (1-t^{\gamma+1}) + \frac{1}{2}\lambda(1-t^2)}{1-\delta}. \quad (11)$$

where $g(\theta_0) = (1-\lambda)\gamma\theta_0^{\gamma-1} + \lambda$. It can be checked that both $\bar{\theta}_0^l(\gamma, \lambda, \delta)$ and $\bar{\theta}_0^h(\gamma, \lambda, \delta)$ are increasing in γ . The same intuition used in the case of mixing applies here.

In the low track, θ_1 lies within the interval $[\underline{l}, \bar{l}]$, where \underline{l} and \bar{l} stand for $\underline{l}(\gamma, \lambda, \delta)$ and $\bar{l}(\gamma, \lambda, \delta)$ and denote the human capital θ_1 acquired in the low track by the “worst” (lowest pre-school achiever) and the “best” (highest pre-school achiever) individuals, respectively:

$$\underline{l}(\gamma, \lambda, \delta) = A(1-\rho)^{\frac{1}{\beta}} \bar{\theta}_0^l. \quad (12)$$

$$\bar{l}(\gamma, \lambda, \delta) = A \left(\rho t^\beta + (1-\rho) \left(\bar{\theta}_0^l \right)^\beta \right)^{\frac{1}{\beta}}. \quad (13)$$

Likewise, in the high track, θ_1 lies within the interval $[\underline{h}, \bar{h}]$, where \underline{h} and \bar{h} stand for $\underline{h}(\gamma, \lambda, \delta)$ and $\bar{h}(\gamma, \lambda, \delta)$ and denote the human capital θ_1 acquired in the high track by the “worst” (lowest pre-school achiever) and the “best” (highest pre-school achiever) individuals, respectively:

$$\underline{h}(\gamma, \lambda, \delta) = A \left(\rho t^\beta + (1-\rho) \left(\bar{\theta}_0^h \right)^\beta \right)^{\frac{1}{\beta}}. \quad (14)$$

$$\bar{h}(\gamma, \lambda, \delta) = A \left(\rho + (1-\rho) \left(\bar{\theta}_0^h \right)^\beta \right)^{\frac{1}{\beta}}. \quad (15)$$

The CDF of θ_1 in a tracking system, denoted by $F_T(\theta_1)$, is:

$$F_T(\theta_1) = \begin{cases} 0 & \text{if } \theta_1 \leq \underline{l} \\ (1 - \lambda)\varphi^l(\theta_1)^\gamma + \lambda\varphi^l(\theta_1) & \text{if } \underline{l} \leq \theta_1 \leq \bar{l} \\ \delta & \text{if } \bar{l} \leq \theta_1 \leq \underline{h} \\ (1 - \lambda)\varphi^h(\theta_1)^\gamma + \lambda\varphi^h(\theta_1) & \text{if } \underline{h} \leq \theta_1 \leq \bar{h} \\ 1 & \text{if } \theta_1 > \bar{h}, \end{cases} \quad (16)$$

where $\varphi^j(\theta_1) = \left(\frac{1}{\rho} \left(\left(\frac{\theta_1}{A} \right)^\beta - (1 - \rho)(\bar{\theta}_0^j)^\beta \right) \right)^{\frac{1}{\beta}}$ for $j = l, h$. The distribution $F_T(\theta_1)$ is (weakly) decreasing in γ . The intuition is similar to the one discussed above for the case of mixing. The expected value of θ_1 in a tracking system is:

$$E_T(\theta_1) = \int_{\underline{l}}^{\bar{l}} \theta_1 f_T(\theta_1) d\theta_1 + \int_{\underline{h}}^{\bar{h}} \theta_1 f_T(\theta_1) d\theta_1 = \bar{h} - \int_{\underline{l}}^{\bar{h}} F_T(\theta_1) d\theta_1, \quad (17)$$

where $f_T(\theta_1)$ denotes the p.d.f. of θ_1 in a tracking system. As in the case of mixing, from Eq. (17) it is immediate that the average human capital in a tracking system increases as the dispersion level in pre-school achievement distribution decreases (measured by an increase in either γ or λ). Again, this is due to the fact that the mean pre-school achievement within both the low and high tracks increases in γ and λ .

3 A comparison of mixing and tracking systems

3.1 A political economy analysis

Suppose first that the educational system is chosen by majority voting and that every individual votes for the system under which her final level of human capital θ_1 would be higher. In this case, exactly a proportion δ of the population would prefer mixing (those with $\theta_0 < t$) because in a tracking system they would be placed in the low track, where they would enjoy a lower peer effect. The rest of the population would prefer tracking (those with $\theta_0 > t$) because they would be placed in the high track, where they would enjoy a higher peer effect. Thus tracking will defeat mixing, under a majority voting rule, whenever $\delta < 1/2$ and mixing will defeat tracking whenever $\delta > 1/2$. Because there is a positive correlation between the probability of being assigned to the high track and family background, the following empirical implication yields: a non-elitist tracking system, i.e., a system in which middle class households get their kids into the high track (that is $\delta < 1/2$), is more likely to be politically supported in a democracy than an elitist tracking system that leaves them out (that is, $\delta > 1/2$).

This result may help to explain how voters' preferences may have shaped the evolution of public perceptions towards tracking systems in the US and the different

patterns of segregation in different European countries. After the Second World War, the majority of European countries reformed their secondary education systems by delaying the point in time at which students are selected into tracks and lines (see [Ariga et al. 2005](#)). One recent example is Spain, where the government's attempt to introduce a tracking system in 2000 was met with a strong social backlash.

3.2 Choosing the education system behind the veil of ignorance

Most works dealing with the effects of segregation in general (e.g., among students or communities) just focus on comparing aggregate or average outcomes in segregated groups, relative to mixed ones. In fact, and as we will see below, this usual approach in the literature can be rationalized as a constrained interpretation of the veil of ignorance approach.

In what follows I will assume that an individual's utility increases with human capital acquired at the compulsory education level, θ_1 . In addition, here I consider that individuals choose the education system behind the so-called "veil of ignorance". That is, they choose between societies without knowing where they will be placed or what characteristics they will have in each society. The idea of choosing from behind a veil of ignorance, to reflect the fairness of societies, has proved very useful in theoretical economics (see the seminal works of [Harsanyi 1953, 1955](#) and [Rawls 1971](#) and more recently [Cremer and Pestieau 1998](#)) and in empirical economics (see [Johansson-Stenman et al. 2002](#) and [Carlsson et al. 2003](#), among others).¹⁴ This is the first study, to our knowledge, that analyses the choice between different education policies from behind the veil of ignorance. This is surprising because it is quite a natural approach in this context. First, observe that real world political debate and specifically, educational debate, is often framed in terms of general principles of justice or of concerns for individual liberty. Thus in this way the veil of ignorance approach can be seen as complementing the standard one. Second, note that families might not be totally certain under which education system their descendants' human capital will be higher. Thus the veil of ignorance approach also incorporates risk and families risk-aversion into the analysis of education policy evaluation. In other words, the veil of ignorance is an adequate approach to evaluate policies that affect agents differently, depending on their position in society, when there is uncertainty about the position each agent would occupy. This is certainly the case for pre-school achievement of the child and track placement in education.

The possible implications of this approach have been intensely debated. According to [Harsanyi \(1953, 1955\)](#), when choosing from behind the veil of ignorance, individuals would choose the alternative that maximizes expected utility or equivalently, they would unanimously agree on the alternative that maximizes a utilitarian welfare function. However, [Rawls \(1971\)](#), who first used the terminology "veil of ignorance", disagreed with the utilitarian proposal and instead argued that each individual would

¹⁴ [Johansson-Stenman et al. \(2002\)](#) and [Carlsson et al. \(2003\)](#) analyze the individuals' choice between alternative societies with different income distributions behind a veil of ignorance. In both experiments they instructed the respondents to consider the well-being of their *imaginary grandchild*, that is, to choose the alternative that would be in the interest of their grandchild.

adopt a “maxi–min” strategy, selecting the alternative that ensures the best of the worst possible outcomes.¹⁵ I will not discuss here which is the most appropriate assumption because, although it is certainly important, this debate falls beyond the scope of this paper.

The Rawlsian approach implies selection of the education system that maximizes the utility of the worst-off individuals in the society.¹⁶ To do this, we have to first define who are the worst-off in our model. If, for example, we take as the worst-off those with pre-school achievement levels below the threshold level and with poor parents, the result is quite immediate. Mixing is the education system chosen. This is derived directly from the properties of the human capital production function because maximizing the utility of these individuals will imply maximization of their human capital at the compulsory level.

Following [Harsanyi \(1953\)](#), individuals behind the veil of ignorance will choose the education system that maximizes expected utility. One possibility is just to compare the two systems in terms of average human capital. This is akin to assuming that all individuals are risk-neutral behind the veil of ignorance. Thus I will analyze the difference between the average human capital under both systems, i.e., $E_T(\theta_1) - E_M(\theta_1)$. It follows from (3) that:

$$\begin{aligned} E_T(\theta_1) - E_M(\theta_1) = & \int_0^t A \left(\rho \theta_0^\beta + (1 - \rho) \left(\bar{\theta}_0^l \right)^\beta \right)^{\frac{1}{\beta}} g(\theta_0) d\theta_0 \\ & + \int_t^1 A \left(\rho \theta_0^\beta + (1 - \rho) \left(\bar{\theta}_0^h \right)^\beta \right)^{\frac{1}{\beta}} g(\theta_0) d\theta_0 \\ & - \int_0^1 A \left(\rho \theta_0^\beta + (1 - \rho) \left(\bar{\theta}_0^m \right)^\beta \right)^{\frac{1}{\beta}} g(\theta_0) d\theta_0, \quad (18) \end{aligned}$$

where $g(\theta_0)$ is defined above. If $\beta = 1$, these integrals can be computed easily, and it can be checked that in this case $E_T(\theta_1) - E_M(\theta_1) = 0$. For $\beta < 1$, I discuss the results using numerical simulations. In addition, I present the results for two ρ values: $\rho = 0.75$ and $\rho = 0.95$.¹⁷ Finally, and in order to check the robustness of the results to the design of tracking system, I present all results for three δ values: $\delta = 1/4$, $\delta = 1/2$ and $\delta = 3/4$.¹⁸

¹⁵ [Johansson-Stenman et al. \(2002\)](#) find empirical evidence that shows that most individuals do not follow the strategy proposed by Rawls, but some do.

¹⁶ Observe that this could be the rationale for some recent education policies in the US, such as the No Child Left Behind Act.

¹⁷ Recent empirical evidence (see, for example, [Heckman \(2006\)](#) and references therein) show that families, not schools, are the major sources of inequalities in school performance, implying that ρ should be high enough.

¹⁸ This means that the threshold level, according to which students are grouped in the low and high tracks, is the first quartile, the median and the third quartile of θ_0 in the population, respectively. Note that a tracking system with $\delta = 0$ or $\delta = 1$ is equivalent to a mixing one.

In particular, I focus here first on the role of the degree of complementarity between peer effects and individuals' pre-school achievement levels (captured by β) and second, on the role of the level of dispersion in the pre-school achievement distribution (captured by γ and λ). To better understand the role of the latter on the comparison between average human capital in the tracking and mixing systems, I analyze it in two societies, A and B, that differ in their initial dispersion of pre-school achievement. This difference might be due to differing pre-school achievement gaps (see Fig. 2a, where I set $\gamma = 3/4$ for society A and $\gamma = 1/4$ for society B) or a different proportion of rich individuals (see Fig. 2b, where I set $\lambda = 3/4$ for society A and $\lambda = 1/4$ for society B). Then, I analyze the impact of variations in the dispersion level of pre-school achievement on the difference between the average human capital under both systems, by examining the effects of changes in λ and γ . Figure 2a represents $E_T(\theta_1) - E_M(\theta_1)$ as a function of λ for different values of β , δ and ρ . Figure 2b represents $E_T(\theta_1) - E_M(\theta_1)$ as a function of γ for different values of β , δ and ρ .¹⁹

Several results are found. First, Fig. 2a and b suggest that $E_T(\theta_1) - E_M(\theta_1)$ is decreasing in ρ . Recall from Eq. (3) that the higher ρ , the lower the weight of peer effects (which is the only determinant of the difference between both systems) on human capital accumulation.

Second, Fig. 2a and b illustrate that if $\beta < 1$, then $E_T(\theta_1) - E_M(\theta_1) > 0$; furthermore, these figures suggest that $E_T(\theta_1) - E_M(\theta_1)$ is decreasing in β . Hence, as $E_T(\theta_1) - E_M(\theta_1) = 0$ when $\beta = 1$, it would even be possible to formally prove that the derivative of $E_T(\theta_1) - E_M(\theta_1)$ with respect to β is negative; unfortunately, this derivative does not seem analytically tractable at first sight. The intuition of this result is as follows. As long as β is not very large, i.e., when $\bar{\theta}_0^j$ and θ_0 have some level of complementarity, then average human capital is always much higher in a tracking system than in a mixing system. When β is large, meaning that the two factors become close substitutes, the difference between average human capital under the two systems disappears. To put it differently, when peer effects matter more for high- (low-) ability students than for low- (high-) ability students, average human capital in a tracking system becomes much larger (closer) than average human capital in a mixing system because this is the system where high- (low-) ability students enjoy a stronger peer effect. Note that this result is in line with previous results in the literature.²⁰

The final lesson we can extract from Fig. 2a and b that is even more important than the previous two is that the initial level of dispersion in the pre-school achievement distribution plays a key role in determining the difference in average human capital in tracking and mixing systems. First, observe that if the initial dispersion level in pre-school achievement is low enough, because either γ or λ are high enough (represented by society A in both figures) and as some sources of dispersion appear (captured by a decrease in either γ or λ), then tracking becomes more and more attractive compared to mixing. Observe that this is true regardless of the threshold level used to separate students in a tracking system and the weight of pre-school achievement on human capital acquisition ρ . However, if the initial dispersion level is not that high, because

¹⁹ In addition, I set $A = 2$ in both Fig. 2a and b.

²⁰ See [Arnott and Rowse \(1987\)](#) and [Benabou \(1996\)](#).

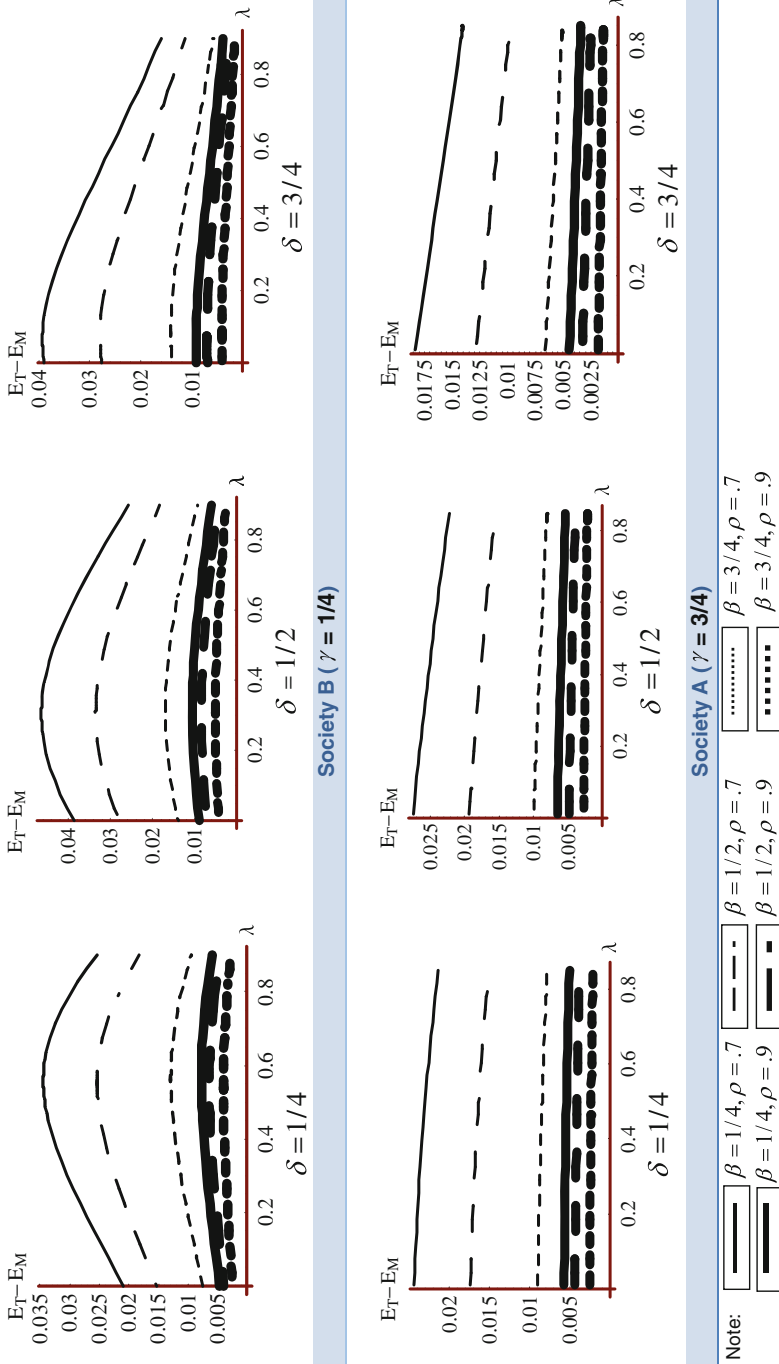


Fig. 2 Difference in average human capital under tracking and mixing

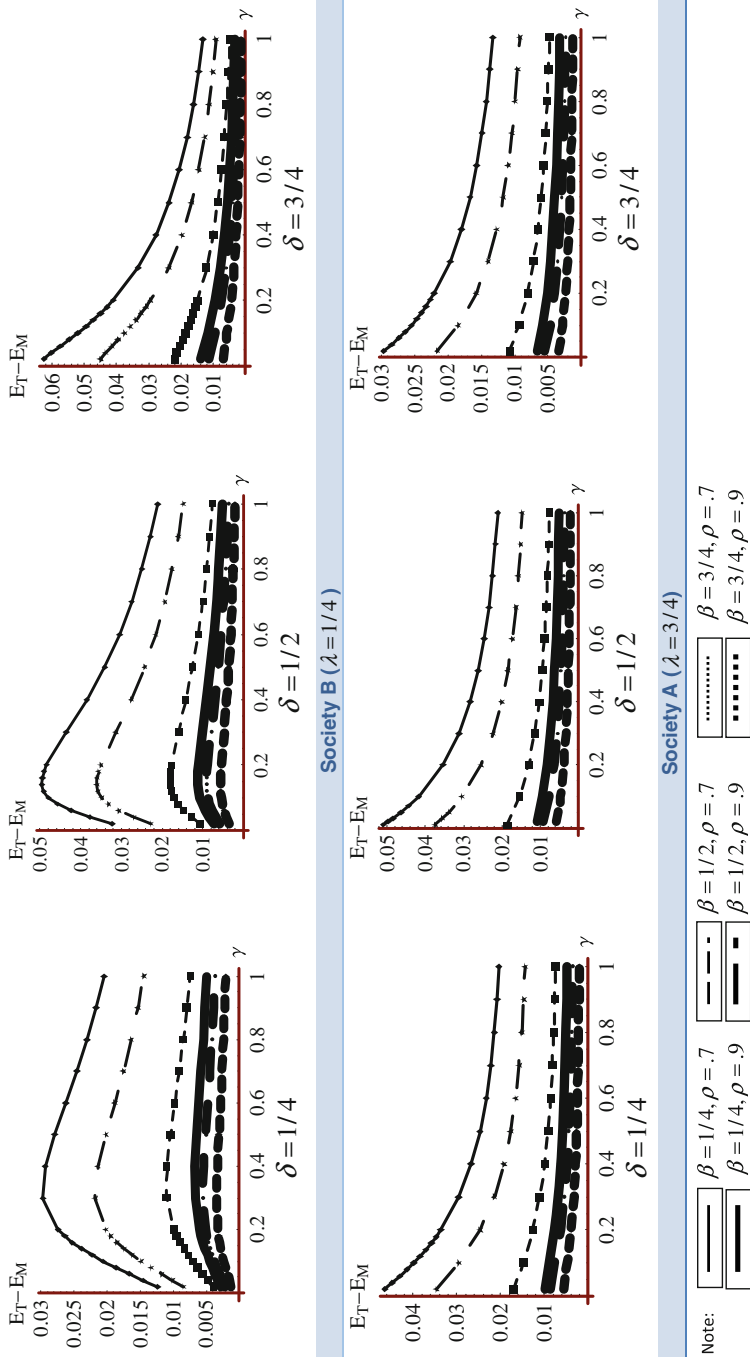


Fig. 2 continued

either γ or λ are very low (see society B in both figures when either γ or λ are very low) then, as the dispersion in pre-school achievement increases (again, measured by a decrease in either γ or λ), tracking becomes less and less attractive compared to mixing. Finally, observe that as the tracking system becomes more elitist (δ increases), and regardless of the initial dispersion level in pre-school achievement, as some sources of dispersion appear, tracking becomes more and more attractive than mixing.

The intuition of the previous result is found in the properties of the education production function. If the initial dispersion level in the pre-school achievement distribution is low enough (because either γ or λ are high enough) and as some dispersion appears (captured by a decrease in either γ or λ), the gain in the high track (as compared to the mixing case) more than compensates for the loss in the low track. The complementarity between the peer group effect and the pre-school achievement level offsets the concavity effect here, and drives the result. The same argument can be applied to the case of an elitist tracking system, where clearly the complementarity effect is exacerbated and makes average human capital in the tracking system much higher than in the mixing system. However, if the initial dispersion level in pre-school achievement distribution is not that low (because either γ or λ are not that high), then further increases in the dispersion (captured again by a decrease in either γ or λ) makes tracking less and less attractive compared to mixing. That is, complementarity between the peer group effect and the pre-school achievement level still offsets the concavity effect, but the losses of those placed in the low track might be balanced with the gains of those placed in the high track (compared to mixing).

This result yields a very clear policy implication. Namely, a government, while implementing policies to reduce the gap in pre-school achievement between rich and poor students, should also choose very carefully the method of grouping them in order to maximize average human capital. In particular, in societies with high initial dispersion levels, and as the government implements policies to reduce the gap in pre-school achievement, tracking will be more and more preferred to mixing. As the dispersion level in pre-school achievement decreases, or in societies with low initial dispersion levels, further efforts to reduce the gap in pre-school achievement make the gains of tracking with respect to mixing diminish.

This result still holds even when comparing two societies with the same mean pre-school achievement but different dispersion levels. In addition, it is robust to the skewness assumption on the pre-school achievement distribution. Both statements can be checked if we assume pre-school achievement to be uniformly distributed on $(0.5 - x, 0.5 + x)$ and reduce x . This reduces the variance but keeps the mean constant. Then regardless of the dispersion level, $E_T(\theta_1) - E_M(\theta_1) = 0$ for $\beta = 1$ and $E_T(\theta_1) - E_M(\theta_1) > 0$ for any $\beta < 1$. Figure 3 shows this result.

It represents $E_T(\theta_1) - E_M(\theta_1)$ as a function of x for different values of β and δ . Regardless of the design of the tracking system δ , the main results found above remain true. First, as peer effects and individual pre-school achievement become close substitutes (β increases), the difference between average human capital under tracking and under mixing systems diminishes. Second, as the level of dispersion in pre-school achievement increases (x increases), the difference between average human capital under both systems increases. This result coincides with the one found in Fig. 2a and b for society A. Finally, observe that using a uniform distribution for pre-school

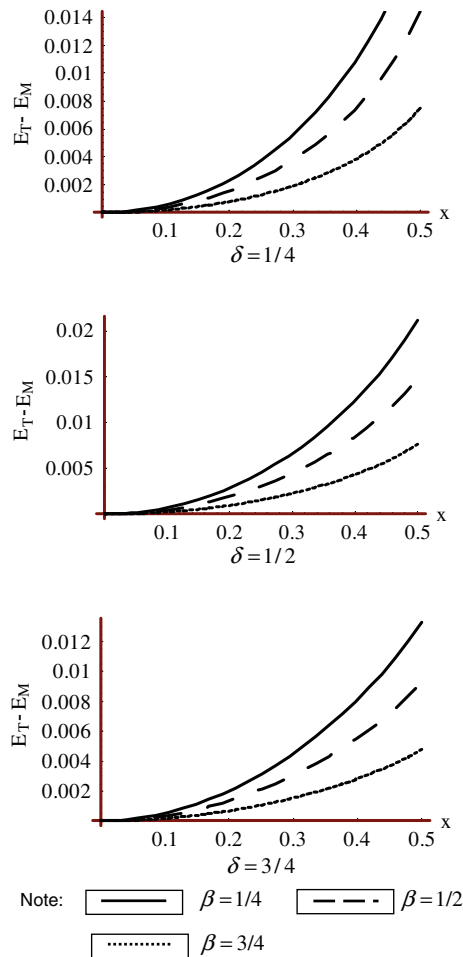


Fig. 3 Difference between average human capital under tracking and mixing. The uniform case

achievement prevents us from considering high dispersion levels in pre-school achievement as the ones considered in Society B in Fig. 2a and b for either low γ or low λ . In that case, there are not only very few rich individuals, but also the mean of pre-school achievement among poor individuals is very low. Therefore, the results above for high initial dispersion levels cannot be replicated here.

Another possibility, while assuming that individuals behave as expected utility maximizers, is to compare the two systems in terms of first order stochastic dominance. Recall that if the distribution of human capital under a given system dominates that of another, according to first order stochastic dominance, all individuals can be said to prefer the former over the latter.²¹ However, in our model neither $F_T(\theta_1)$ first-order stochastically dominates $F_M(\theta_1)$ (because $\underline{l} < \underline{m}$ and hence $F_M(\theta_1) - F_T(\theta_1) < 0$ if

²¹ See the seminal paper by Rothschild and Stiglitz (1970).

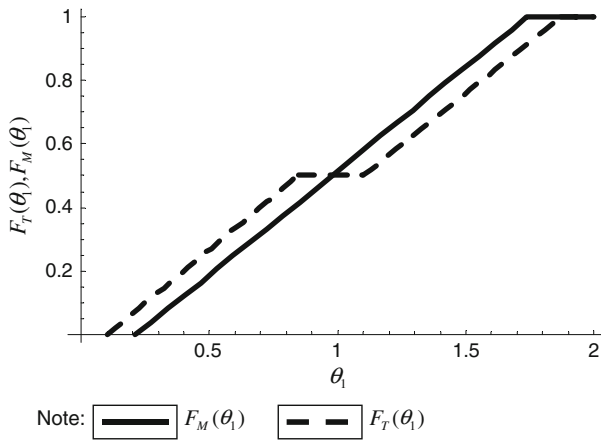


Fig. 4 No first order stochastic dominance

$\underline{l} < \theta_1 < \underline{m}$), nor $F_M(\theta_1)$ first-order stochastically dominates $F_T(\theta_1)$ (because $\bar{m} < \bar{h}$ and hence $F_T(\theta_1) - F_M(\theta_1) < 0$ if $\bar{m} < \theta_1 < \bar{h}$). Figure 4 illustrates this result, where $F_M(\theta_1)$ and $F_T(\theta_1)$ are represented in solid and dashed lines, respectively.²²

Finally, I will consider that all individuals behind the “veil of ignorance” are risk-averse. In this case, they will prefer the less risky distribution of human capital. This criterion leads to the concept of second order stochastic dominance (SOSD). But, once again, in general there is no preferred system according to this criterion. On the one hand, $F_T(\theta_1)$ does not second-order dominate $F_M(\theta_1)$ because $\int_0^{\underline{m}} (F_T(\theta_1) - F_M(\theta_1))d\theta_1 = -\int_{\underline{l}}^{\underline{m}} F_T(\theta_1)d\theta_1 < 0$. On the other hand, using (8) and (17), it follows that $\int_0^{\bar{h}} (F_T(\theta_1) - F_M(\theta_1))d\theta_1 = \int_{\underline{l}}^{\bar{h}} F_T(\theta_1)d\theta_1 - \left(\int_{\underline{m}}^{\bar{m}} F_M(\theta_1)d\theta_1 + \bar{h} - \bar{m} \right) = -(E_T(\theta_1) - E_M(\theta_1))$; as Fig. 2a and b illustrate, if $\beta < 1$ then $E_T(\theta_1) - E_M(\theta_1) > 0$ and hence $\int_0^{\bar{h}} (F_T(\theta_1) - F_M(\theta_1))d\theta_1 < 0$, i.e., $F_M(\theta_1)$ does not second-order dominate $F_T(\theta_1)$.

In other words, in a situation in which peer effects matter more for high achievers than low achievers (that is, if β is low), then the distribution of human capital in a tracking system is more “spread” than in a situation in which peer effects matter more for low achievers than high achievers. Hence in this case the distribution of human capital in a tracking system cannot be considered less risky than in a mixing system. Thus risk-averse individuals will not prefer tracking rather than mixing systems. However, because average human capital is always higher in tracking rather than mixing systems, individuals will not prefer mixing over tracking systems. Similarly, as β increases,

²² The following parameter values are assumed here: $A = 2$, $\delta = 1/2$, $\gamma = 0.8$, $\lambda = 5/6$, $\rho = 3/4$ and $\beta = 0.9$.

tracking can be considered less and less risky than mixing. However, the difference between average human capital under both systems diminishes. As a result, among all risk-averse households with uncertainty about their kids' pre-school achievements, there will be no unanimity in the choice of the educational system. Alternatively, we can consider that a social planner with SOSD as a justice criterion would not prefer mixing nor tracking.

The general message we can extract is that tracking is the system that maximizes average human capital, although, given the choice between tracking and mixing, there would be no unanimously preferred system for the population. In addition, it is found that in societies where pre-school achievement is not very dispersed, as some sources of dispersion in pre-school achievement appear, then tracking becomes more and more attractive than mixing for maximizing average human capital. However, this is not the case in societies where pre-school achievement is much dispersed. There, additional dispersion makes tracking relatively less attractive than mixing for maximizing average human capital.²³

3.3 On the robustness of the results

A crucial aspect of the analysis is the choice of the specific Education Production Function (EPF) in Eq. (3). This is justified because it allows the comparison of the two systems with numerical examples when it is not possible to obtain closed-form analytical solutions. Nonetheless, restricting attention to a specific production function imposes generality losses. In this case, the specification of peer effects in (3) does not account for the impact of heterogeneity, as suggested by the empirical evidence found by [Hoxby and Weingarth \(2006\)](#). In particular, they find that the peer effect depends on students' characteristics, and they provide some support for a specification in which every student learns the most when she is with students like her (see also [Manski and Wise 1983](#)). Now, suppose we consider an alternative specification for the EPF that captures this empirical evidence. Notice that this pro-homogeneity specification would underlie support for tracking because not only do good students gain, compared to mixing, but weak students might as well. In particular, weak students may be better off in a tracking system if the gain from being with peers slightly brighter than them compensates for the losses derived from a lower mean pre-school achievement within the group than it would be in a mixing system. Therefore, would the results of the paper still maintain? The answer is positive in most cases. Specifically, the main results of the paper will maintain. Recall that even when ruling out the evidence found by [Hoxby and Weingarth \(2006\)](#), I find that tracking is the system that maximizes average human capital at the end of the compulsory level. Thus such a specification capturing the [Hoxby and Weingarth \(2006\)](#) evidence would only reinforce this finding. However, the result regarding the education system chosen under the Rawlsian approach might be reversed now. In this case, individuals with pre-school achievement below

²³ This result contrasts to that of [Kremer and Maskin \(1996\)](#), who find that increases in skill dispersion *always* promote segregation of workers by skill. However, note that they do not consider the existence of peer effects, which are a crucial input for explaining the role of skill dispersion on the optimal allocation of individuals with different skill levels.

the threshold t , and with poor parents (i.e., the worst-off individuals in the society), might be better off in a tracking system. In addition, the result regarding the comparison of the two systems in terms of first order stochastic dominance might not be true now. In this case, and as long as individuals with pre-school achievement below the threshold t are better-off in a tracking system than in a mixing system, then the former might dominate the latter, according to this criterion.

Another related weakness of the EPF in (3) is that it is assumed to be the same in both tracked and mixed classrooms, although, in reality, it might not always be the case. For example, some proponents of tracking argue that by narrowing the range of individuals' abilities, teachers can target instruction to a level more closely aligned with other students' needs than would be possible in more heterogenous environments, implying that the process of acquisition of human capital differs under tracking and mixing systems. Therefore, removing the specification proposed in this paper will only reinforce my main results without adding further insights.

4 Concluding remarks

In this paper, I have analyzed public intervention in education when the government, taking into account the existence of peer effects, has to decide how to group students. I have considered two different education systems: tracking and mixing. A number of previous works have studied the optimal education system by focusing on mean achievement. This paper contributes to this line of research by recognizing the degree of dispersion in the pre-school achievement distribution to be a key variable that modulates the impact of non-linearities and crucially affects the comparison between tracking and mixing systems.

The main result of the paper is that tracking is the educational system that maximizes average human capital, regardless of the initial dispersion in the pre-school achievement distribution. However, the gains of tracking compared to mixing are sensitive to changes in the level of dispersion of pre-school achievement. In particular, in societies with high initial dispersion levels, and as the government implements policies to reduce the gap in pre-school achievement, tracking will be more and more preferred to mixing. As the dispersion level in pre-school achievement decreases in these societies, or in societies with low initial dispersion levels, further decreases in dispersion levels imply that the gains of tracking with respect to mixing diminish.

This paper allows for some extensions. An important one is the introduction of prices, which are omitted in this paper under the assumption of free education in both systems. Finally, I think that the results presented here are relevant to several recent debates in the literature on the economics of education. There is increasing evidence that shows the early emergence and persistence of gaps in cognitive and non-cognitive skills (see, among others, [Carneiro and Heckman 2003](#)). Studies that highlight the importance of increasing expenditures on early childhood care in pursuing both efficiency and equity provide an interesting illustration. As I have showed in this paper, a government, while reducing the gap in pre-school achievement between rich and poor students, should also choose very carefully the method of grouping them, in order to maximize average human capital or to find the preferred educational system

in the population. Another example is the literature that looks at the heterogeneity in grouping policies across countries and tries to explain it (see [Brunello et al. 2005](#); [Ariga et al. 2005](#), among others). As [Brunello et al. \(2005\)](#) pointed out, efficiency considerations are insufficient to explain the existing differences that instead might be driven by some distributional concerns of society.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Angrist JD, Lang K (2004) Does school integration generate peer effects? Evidence from Boston's Metco program. *Am Econ Rev* 94(5):1613–1634
- Ariga K, Brunello G, Iwahashi R, Rocco L (2005) Why is the timing of school tracking so heterogeneous? IZA discussion paper no. 1854
- Arnott RJ, Rowse J (1987) Peer group effects and educational attainment. *J Public Econ* 32:237–306
- Becker GS (1993) Human capital: a theoretical and empirical analysis with special reference to education. The University of Chicago Press, Chicago
- Benabou R (1996) Efficiency and equity in human capital investment: the local connection. *Rev Econ Stud* 63(2):237–264
- Bishop J (2006) Drinking from the fountain of knowledge: student incentive to study and learn externalities, information problems and peer pressure. In: Hanushek EA, Welch F (eds) *Handbook of economics of education*, vol 2. Elsevier, Amsterdam
- Brunello G, Giannini M, Ariga K (2005) The optimal timing of school tracking. In: Peterson P, Woessmann L (eds) *Schools and the equal opportunity problem*, MIT Press, Cambridge
- Brunello G, Checchi D (2007) Does school tracking affect equality of opportunity? New international evidence. *Econ Policy* 22(52):781–861
- Carlsson F, Gupta G, Johansson-Stenman O (2003) Choosing from behind a veil of ignorance in India. *Appl Econ Lett* 10:825–827
- Carneiro P, Heckman JJ (2003) Inequality in America: What role for human capital policies? In: Heckman JJ, Krueger AB, Friedman B, (eds) Chapter 2, MIT Press, Cambridge, pp 77–237
- Collado D, Iturbe-Ormaetxe I (2008) Public transfer to the poor: is really Europe more generous than the United States. IVIE Working Paper 2008-05, forthcoming *International Tax and Public Finance*
- Cremer H, Pestieau P (1998) Social insurance, majority voting and labor mobility. *J Public Econ* 68(3): 397–420
- Ding W, Lehrer SF (2007) Do peers affect student achievement in China's secondary schools? *Rev Econ Stat* 89(2):300–312
- Eppl D, Romano RE (1998) Competition between private and public schools, vouchers and peer group effects. *Am Econ Rev* 88(1):33–62
- Eppl D, Newlon E, Romano RE (2002) Ability tracking, school competition, and the distribution of educational benefits. *J Public Econ* 83:1–48
- Foster G (2006) It's not your peers, and it's not your friends: some progress toward understanding the educational peer effect mechanism. *J Public Econ* 90:1455–1475
- Gradstein M, Justman M (1997) Democratic choice of an education system: implications for growth and income distribution. *J Econ Growth* 2:169–183
- Harsanyi J (1953) Cardinal utility in welfare economics and the theory of risk taking. *J Polit Econ* 61: 434–435
- Harsanyi J (1955) Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *J Polit Econ* 63:309–321
- Heckman JJ (2006) Skill formation and the economics of investing in disadvantaged children. *Science* 312(5782):1900–1902
- Hoxby CM (2000) Peer effects in the classroom: learning from gender and race variation. NBER WP No. 7867

- Hoxby CM, Weingarth G (2006) Taking race out of the equation: school reassignment and the structure of peer effects. Harvard University manuscript
- Johansson-Stenman O, Carlsson F, Daruvala D (2002) Measuring future grandparents' preferences for equality and relative standing. *Econ J* 112:362–383
- Kang C (2007) Classroom peer effects and academic achievement: quasi randomization evidence from South Korea. *J Urban Econ* 61:458–495
- Kremer M, Maskin E (1996) Wage inequality and segregation by skill. NBER Working Papers No. 5718, forthcoming. *Q J Econ*
- Manski CF, Wise DA (1983) College choice in America. Harvard University Press, Cambridge
- OECD (2002) OECD economic studies, No. 34, 2002/01
- OECD (2004) Learning for tomorrow: first results from PISA 2003, OECD, Paris
- Rawls J (1971) A theory of justice. Harvard University Press, Cambridge
- Rothschild M, Stiglitz J (1970) Increasing risk: I A definition. *J Econ Theory* 2(3):225–243